

Visualizing and Exploring Data

Data Visualizing

Data visualization is the graphical representation of datasets and information. Data visualization is an umbrella term for visualizing all types of data through charts, graphs, and maps.

The ultimate goal is to visually represent your data in an accessible and easy-to-understand manner. Visualizing data is a fundamental step in understanding trends, uncovering patterns, and tracking outliers.

Types of Data Visualization

1. Column Chart

They are a straightforward, time-tested method of comparing several collections of data. A column chart may be used to track data sets across time.

2. Line Graph

A line graph is used to show trends, development, or changes through time. As a result, it functions best when your data collection is continuous as opposed to having many beginnings and ends.

3. Pie Chart

In a pie chart, a single, constant number is represented by the several categories that make up its parts. You will portray numerical quantities in percentages when you employ one. All of the various components should sum up to a hundred percent when totaled.

4. Bar Chart

To compare data along two axes, use bar charts. A visual representation of the categories or subjects being measured is shown on one of the axes, which is numerical.

5. Heat Maps

A data visualization method that uses colors to denote values; great for seeing trends in huge datasets.

6. Scatter Plot

The correlation between variables is examined using a scatter plot. At the point where the data's two values overlap, the data are represented on the graph as dots.

7. Bubble Chart

A variant of the scatter plot where the size and color of the bubbles, which represent the data points, provide extra information, are used to depict the data points as dots.

8. Funnel Chart

To illustrate a sequential process from top to bottom, a funnel chart's principal purpose is to represent it graphically. As the process flows down, the amount generally decreases, making the data set at the top of the process greater than the bottom.

9. Radar Chart

Radar charts are a sort of data visualization that aids in the analysis of objects or categories in light of a variety of attributes. The radar chart consists of a circle with concentric rings, and the data are shown as dots on the chart. The shape is then formed by connecting the dots. Each thing or group has a shape.

10. Tree Chart

An alternative to a table for precise numerical data is a tree chart, often known as a tree diagram. The basic goal of a tree chart is to represent data as pieces of a larger whole within a category.

11. Flow Chart

One extremely adaptable method of data display is the flowchart. Use mind maps for brainstorming, flowcharts to depict a process graphically and hierarchical data of objects or people.

12. Gauge

A gauge is a percentage visualization. There are a few uses for the half-doughnut-like form. To display a percentage figure with an arrow pointing to it is the simplest use. If you have a small quantity of data to work with, this is a fantastic option.

13. Gantt Chart

Horizontal bar graphs are the basis for the Gantt chart; however, they differ significantly from them. A rectangle that extends from left to right stands for each item on the chart. Depending on how long each activity takes to accomplish, each one varies in size.

14. Venn Diagram

A Venn diagram is a data visualization that compares two or more objects by emphasizing their similarities. The most typical Venn diagram design consists of two overlapping circles.

15. Histogram

While a histogram and a bar graph are similar, they use distinct charting systems. The ideal sort of data visualization for frequency-based analysis of data ranges is a histogram.

16. Waterfall Chart

A style of bar graph that demonstrates how a sequence of positive and negative numbers affects an initial value.

17. Marimekko Chart

A graphic depiction called a Marimekko chart shows category data using stacked bar graphs of various widths. Mekko charts and mosaic plots are other names for the same type of diagram.

18. Choropleth Map

The technique of color mapping symbology is used to create choropleth maps, which are themed maps used to display statistical data. It shows geographically segmented sections or regions that are colored, shaded, or patterned according to a data variable, known as enumeration units.

19. PERT Chart

PERT is a technique for calculating the least amount of time needed to finish a project by analyzing the amount of time needed to complete each job and the dependencies related to it.

20. Dichotomous Key

An identification chart called a dichotomous key allows users to choose between questions and assertions offered in the chart to arrive at a conclusion that will assist them in identifying objects or anything else.

21. Mind Map

Using a radial layout to represent thoughts and ideas, mind maps are data visualization that helps organize and spark ideas while dealing with complicated material.

22. Timeline

They are visual depictions of a historical period with significant events labeled in chronological order. They may be more detailed visuals or rather straightforward linear representations.

23. Concentric Circles

A style of data visualization known as concentric circles makes use of circles inside circles to represent hierarchical connections or proportions, with the size of the circles signifying the amount of data being displayed.

24. Radial Wheel

With each spoke or segment denoting a separate category or value, radial wheels are a style of data visualization that uses a circular structure to highlight connections between data elements.

25. Percentage Bar

A type of data visualization known as percentage bars use a horizontal bar with proportional segments to show numbers as percentages of the total and the relative size of each category.

26. Donut Chart

Donut charts, often called doughnut charts, are variants on pie charts that include a hole in the center, giving them the appearance of doughnuts. This open space may be used to display further information.

27. Half-Donut Chart

The half-doughnut chart is precisely what its name suggests—it's a half-doughnut chart. When displaying modest amounts of data, this type of data visualization is a useful option. A half-donut chart should, ideally, not include more than three wedges.

28. Polar Graph

If the data values are substantially dissimilar from one another, choose a polar graph as the types of data visualization in data science. If not, it could be difficult to read at a glance.

29. Icon Array

Icon arrays are a type of data visualization that works well for displaying proportions and patterns because they employ icons or symbols to represent individual data points, such as circles or squares.

30. Cone Chart

Hierarchy is depicted with a cone chart. The greatest value data is located on the broadest section of the cone. The other values are distributed in descending order from top to bottom of the cone.

Data Exploration

Data exploration is the first step in data analysis involving the use of data visualization tools and statistical techniques to uncover data set characteristics and initial patterns.

During exploration, raw data is typically reviewed with a combination of manual workflows and automated data exploration techniques to visually explore data sets; look for similarities, patterns and outliers; and identify the relationships between different variables.

Data exploration is also sometimes referred to as *exploratory data analysis*, which is a statistical technique used to analyze raw data sets in search of their broad characteristics.

Why is data exploration important?

Humans are visual learners, able to process visual data more easily than numerical data. Consequently, it's challenging for data scientists to review thousands of rows of data points and infer meaning without assistance.

Data visualization tools and elements such as colors, shapes, lines, graphs and angles aid in effective data exploration of metadata, enabling relationships or anomalies to be detected.

How does data exploration work?

There are three general steps included in data explanation:

1. **Understand your variables.** To explore a data set meaningfully, it's essential to know the nature of all the data in it. This is usually clear from the names and the descriptions of the data columns, the data types and other associated metadata. These can usually be found in an organization's data catalog.
2. **Search for outliers.** A data set can include outliers -- data points that stray far from the average of similar data points. They can distort the exploration of the data by skewing averages and other statistics that emerge during exploration. Such outliers can be found by

visually plotting the data in scatterplots, histograms or other graphics. Once identified, data scientists can remove, ignore or investigate outliers as needed.

3. **Look for patterns and relationships between data items.** Visual data plots as well as a broad range of analytics tools can reveal patterns within data sets that provide additional layers of meaning and insights. This, in fact, is the primary product of data exploration: to discover value in data that wasn't apparent before. These insights can lead to using the data to explain past events, facilitate predictive modeling and select the best courses of action.

What industries use data exploration?

Any business or industry that collects or uses data can benefit from data exploration. In fact, it's difficult to conceive of an industry that wouldn't. Some of the more prominent industries where data exploration is prevalent include the following:

- Software development.
- Engineering.
- Healthcare and medicine.
- Education.

